1.2.4.1   The General Format of an Answer to a Query Follows:

To: (person or number of person making request)

Category: (000 to 255; to search all categories one would
                       use 000. The category number refers to a list
                       of categories that would be used to subdivide
                       the file, i.e. 002 might be Cancer-tobacco use)

Descriptors: (terms used by the user to search for this
                       document; this list of terms or single term
                       would also be associated with the weight number(s)
                       and the threshold value ascribed to the search
                       that would determine retrieval items)

Accession numbers: (accession numbers will be presented to enable
                       the user to find the document in the aperature
                       card file. The accession numbers will be unique
                       for documents and will be designed to show the
                       year of acceptance into the system as well as the
                       sequential number given to the document in that
                       year, i.e. 6600024 The first two digits refer
                       to the year and the next five digits refer to
                       the number of the document for that year.)

1.3   Continual Operation of the IS&R System

The entire IS&R system is designed so that all rules and regulations for
abstracting, scanning, searching, indexing, etc., are written and therefore the
operation is not dependent upon any person or even a company. Operation of the
entire system can be continued for Covington & Burling or for other Tobacco Industry
lawyers by other companies, if it is so desired at some future date.

1.4   General Examples of Typical Questions that could be asked by the User
                       "Give me all articles written by P. Q. Smith
                       only when he was the senior author, and when
                       he wrote about paper filters and lung cancer."

In the above example the keys to the articles are underlined. All in-
formation for documents that were indexed using all of these terms would be
retrieved.

2010008079

1-5

3i

If the number of articles retrieved was too great to handle, further specificity could be achieved by specifying, e.g., "since 1963" or a specific type of paper filter or cancer.

"What was the death rate for emphysema in 1962?"

This is a request for a specific piece of numerical information. Although the system is a document retrieval system (one that tells you where you can find the answer) rather than a data retrieval system, it is relatively easy to locate the document that has the answer in it. Using the underlined words above, you would be referred to a document, or a page of a document, and the death rate could be easily obtained from the microfilm reader.

1.4.1 Example of the Processing of a Multi-part Document such as a Book or the Published Proceedings of a Meeting

Cigarette Labeling and Advertising: Hearings before the Committee on Commerce, United States Senate, Eighty-Ninth Congress, First Session on S. 559 and S. 547, bills to regulate labeling of cigarettes and for other purposes. Part I: March 22, 23, 25, 29, 30, April 1 and 2, 1965.

For purposes of accessioning, indexing, annotating, and eventual retrieval, this 1625-page book would be broken up into the various contributions, many but not all of which are listed in the Table of Contents, of which it is composed. These separate documents which would be accessioned would include the texts of the two bills, the address by Luther Terry from the Congressional Record, the comments by Senator Bennett and Senator Moss, the Agency comments referred to in the Table of Contents, the 46 statements listed, with the discussion following each, the 30 submitted statements, and the various letters, wires, and resolutions which are printed in the last section of the book.

Samples of other questions and answers are attached as an Exhibit.

1.5 Why the Machine Search System has been Selected over the Manual Method

Machine searchable coordinate indexes are necessary when any or a combination of the following characteristics is required: deep indexing (an average of 10 terms or more is considered "deep"; 50 term average indexing for the tobacco industry is proposed), relatively large collections (over 15,000 items; this number will be exceeded after the first year of operation), logical capability (the ability to perform logical unions, intersections, negations, differences, of descriptors, etc.), fast retrieval, and high degree of completeness.